

Instituto de Matemática e Estatística da Universidade
de São Paulo - IME-USP

Projeto de Iniciação Científica

Análise de Dados Circulares

Aluno:

Rafael Izbicki

Orientador:

Prof. Dr. Luís Gustavo Esteves

10 de Dezembro de 2008

Sumário

1	Introdução	2
2	Resumo	2
3	Métodos e Resultados	3
3.1	Notações e Convenções	3
3.2	Análise Descritiva	4
3.3	Modelos Probabilísticos Circulares	7
3.4	Inferência	11
3.4.1	Distâncias e Funções de Perda no Círculo	11
3.4.2	Resultados	13
4	Conclusões	22
A	Demonstrações adicionais	24
A.1	Teorema (3.2)	24
A.2	Teorema (3.5)	25

1 Introdução

Medidas circulares ocorrem naturalmente em muitas áreas do conhecimento, tais como biologia, geografia e meteorologia, dentre outras (Mardia, [6]). Em geral, medidas circulares são registradas em fenômenos com observações direcionais (direção do movimento de um animal após um determinado estímulo, direção do vento, etc.), assim como em fenômenos periódicos (horário de chegada de um paciente a um hospital).

Em decorrência da natureza periódica de tais características, as técnicas estatísticas desenvolvidas para variáveis reais (na reta) não são, em geral, apropriadas (Fisher, [4]). Deste modo, algumas técnicas de descrição de dados e procedimentos de inferência estatística específicos foram desenvolvidos durante as últimas décadas para situações envolvendo características circulares.

Nesse projeto, pretendemos fazer um estudo do problema de análise de dados circulares, englobando as principais técnicas de descrição de dados, os modelos probabilísticos usualmente adotados e os procedimentos de inferência estatística desenvolvidos, sob as abordagens clássica e bayesiana.

2 Resumo

Neste projeto, vários aspectos relacionados a medidas circulares foram estudados.

Como técnicas de análise para dados na reta mostram-se inadequadas para o tratamento de dados circulares, novas medidas de tendência central, assim como medidas de variabilidade, precisam ser definidas. Desta forma, inicialmente, estudaram-se algumas destas medidas, bem como a sua utilização em análises descritivas de dados circulares (seção (3.2)).

Posteriormente, alguns modelos probabilísticos usualmente adotados para dados circulares foram estudados. Mais especificamente, estudou-se a distribuição uniforme e a distribuição de von Mises. Também foi estudada uma forma de se caracterizar a distribuição de von Mises (seção (3.3)).

Foram também contemplados alguns aspectos relativos à inferência para dados circulares, tanto sob a abordagem clássica quanto sob a abordagem bayesiana, quando consideramos uma amostra aleatória simples da distribuição de von Mises (seção (3.4)), como estimação pontual e testes de hipótese. Em particular, estudou-se o teste para uma hipótese precisa (ou seja, com medida de Lebesgue nula (Pereira e Stern, [12])) e aplicou-se o FBST a este caso. Algumas propriedades deste teste foram então analisadas.

Algumas convenções usuais utilizadas no estudo de dados circulares encontram-se na seção (3.1).

3 Métodos e Resultados

3.1 Notações e Convenções

Para o estudo de dados circulares, algumas convenções são em geral feitas. Primeiramente, variáveis aleatórias provindas de amostras são representadas por letras gregas, assim como parâmetros da população de interesse. Isso não é usual na estatística para dados na reta. Além disso, não é costume diferenciar variáveis aleatórias ainda não observadas (grafadas usualmente em letras maiúsculas) de seus valores já observados (grafados usualmente em letras minúsculas).

Cada medida circular pode ser vista como um ponto em um círculo de raio 1. Uma amostra de n dados circulares é usualmente representada por ϕ_1, \dots, ϕ_n , com $0 \leq \phi_i < 2\pi$, $\forall i \in \{1, \dots, n\}$, em que ϕ_i é o ângulo entre o eixo x e o i -ésimo ponto observado, no sentido anti-horário, como mostrado na Figura 1.

Uma notação comumente empregada é a representação de pontos em um plano através de números complexos. Para tanto, é costume utilizar a fórmula de Euler, $e^{i\phi} = \cos \phi + i \sin \phi$. Um número $(a, b) \in \mathbb{R}^2$ sempre pode ser representado por suas coordenadas polares, (ρ, ϕ) , que satisfazem a relação $\rho e^{i\phi} = a + ib$. Observa-se que ρ é o comprimento de (a, b) e ϕ é o ângulo entre o eixo x e esse vetor, no sentido anti-horário. Uma representação encontra-se

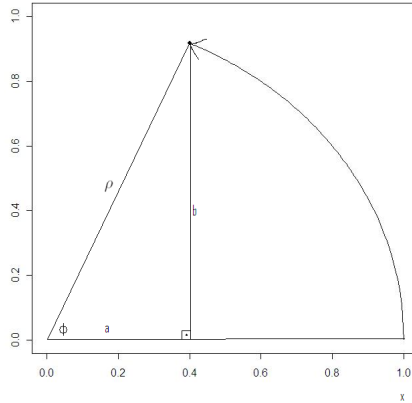


Figura 1: Representação de um dado circular

na Figura 1.

Uma família de funções de grande utilidade no presente estudo é formada pelas funções de Bessel modificadas do primeiro tipo (Mardia et al, [9]), i.e.,

$$I_p(\kappa) = \frac{1}{2\pi} \int_0^{2\pi} \cos(p\phi) \exp[\kappa \cos \phi] d\phi, \quad p \in \{0, 1, \dots\}.$$

Diz-se que $I_p(\kappa)$ é a função de Bessel modificada do primeiro tipo de ordem p aplicada no ponto κ . Alguns modelos probabilísticos envolvem tais funções.

3.2 Análise Descritiva

Com o objetivo de representar as observações graficamente, usualmente utiliza-se o gráfico de dispersão no círculo, em que cada observação é representada por um ponto. Um exemplo deste gráfico encontra-se na Figura 2, para um conjunto de dados simulados. Alternativamente, pode-se usar também o chamado *rose diagram* (Mardia et al, [9]), como mostrado na Figura 3, para o mesmo conjunto de dados. A área de cada setor desse gráfico é proporcional à frequência observada na amostra dos valores pertencentes à respectiva região, como em um histograma usual. Esses gráficos podem ser facilmente contruídos através de funções dos pacotes *CircStats* e *circular* (Jammalamadaka et al, [5]), implementados para o R [13].

Para dados circulares, não faz sentido utilizar a média amostral, adotada usualmente para dados na reta, como medida de centralidade. Por exemplo, suponhamos que estamos

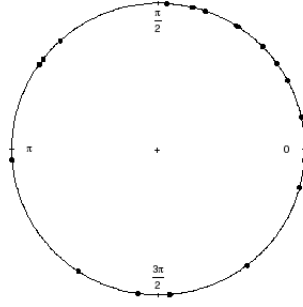


Figura 2: Diagrama de dispersão para dados circulares

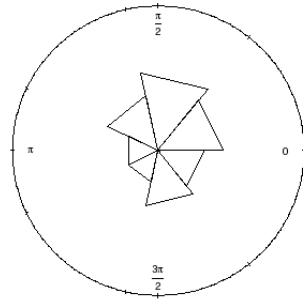


Figura 3: Gráfico do tipo *rose diagram*

interessados em saber o horário em que costumam ocorrer assaltos em um determinado local e temos uma amostra de tamanho 2 com os valores (23:00hs, 01:00hs). Então, convertendo essa amostra para ângulos e calculando a média amostral, obtém-se o equivalente a 12:00hs, valor que evidentemente é inadequado.

Deste modo, dada uma amostra ϕ_1, \dots, ϕ_n , define-se como direção média (ou ângulo médio), $\bar{\phi}$, o ângulo entre o eixo x e o vetor $(\frac{\sum_{i=1}^n \cos \phi_i}{n}, \frac{\sum_{i=1}^n \sin \phi_i}{n}) \equiv (\bar{C}, \bar{S})$. Tal vetor é o centro de massa dos pontos $(\cos \phi_1, \sin \phi_1), \dots, (\cos \phi_n, \sin \phi_n)$ em \mathfrak{R}^2 . O comprimento de (\bar{C}, \bar{S}) , $(\bar{C}^2 + \bar{S}^2)^{1/2} \equiv \bar{R}$, chamado de comprimento médio resultante, é utilizado como medida de concentração dos dados. Se os ângulos observados concentram-se em torno do ângulo médio $\bar{\phi}$, \bar{R} tende a ser maior que quando eles estão mais dispersos. A Figura 4 ilustra estas situações. Observa-se que $0 \leq \bar{R} \leq 1$ e que $\bar{R}e^{i\bar{\phi}} = \bar{C} + i\bar{S}$.

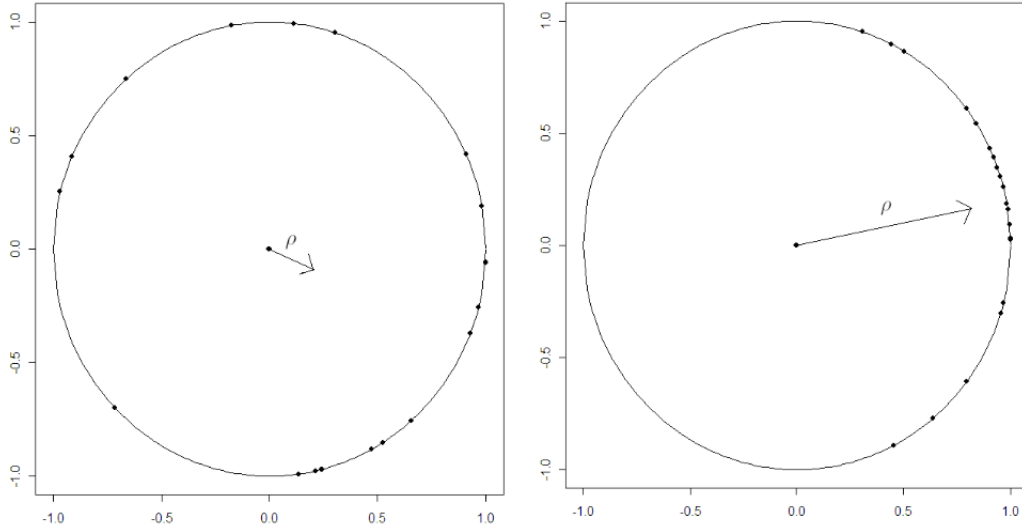


Figura 4: Comparação entre dois ρ 's diferentes

A partir de \bar{R} , podem-se criar diferentes medidas de dispersão, apesar de sua utilização não ser usual. Um medida simples de dispersão é dada por $V = 1 - \bar{R}$, chamada de variância circular amostral. É importante ressaltar que $0 \leq V \leq 1$. Usualmente, define-se o desvio padrão circular como $v = [-2 \log(1 - V)]^{1/2}$. Tem-se que $v \geq 0$. Maiores detalhes sobre tais medidas podem ser encontrados em (Mardia et al, [9]).

Uma variável aleatória circular ϕ é uma variável aleatória cujo suporte é $[0, 2\pi)$. Para uma variável aleatória circular ϕ com função distribuição de probabilidades $F(\phi)$, define-se μ_ϕ , a direção média (ou ângulo médio), e ρ_ϕ , o comprimento médio resultante (Mardia et al, [9]), através da relação

$$\rho_\phi e^{i\mu_\phi} = \int_0^{2\pi} \exp[i\phi] dF(\phi) = \mathbb{E}[\cos \phi] + i\mathbb{E}[\sin \phi]. \quad (3.2.1)$$

A variância circular, ν_ϕ , e o desvio-padrão circular, σ_ϕ , são definidos como $\nu_\phi = 1 - \rho_\phi$ e $\sigma_\phi = [-2 \log(1 - \nu_\phi)]^{1/2}$. A partir de agora, sempre que possível, o índice ϕ será omitido com a finalidade de não carregar a notação.

As seguintes relações serão usadas mais adiante:

$$\sum_{i=1}^n \cos(\phi_i - \mu) = \sum_{i=1}^n (\cos \phi_i \cos \mu + \sin \phi_i \sin \mu) = \cos \mu \left(\sum_{i=1}^n \cos \phi_i \right) + \sin \mu \left(\sum_{i=1}^n \sin \phi_i \right) =$$

$$= n\bar{R}[\cos \mu \cos \bar{\phi} + \text{sen } \mu \text{sen } \bar{\phi}] = n\bar{R} \cos(\bar{\phi} - \mu) \quad (3.2.2)$$

e

$$\sum_{i=1}^n \text{sen}(\phi_i - \bar{\phi}) = \sum_{i=1}^n (\text{sen } \phi_i \cos \bar{\phi} - \cos \phi_i \text{sen } \bar{\phi}) = n\bar{S} \cos \bar{\phi} - n\bar{C} \text{sen } \bar{\phi} = 0. \quad (3.2.3)$$

3.3 Modelos Probabilísticos Circulares

A seguir apresentamos duas distribuições de probabilidade circulares importantes.

Definição 3.1. *Seja ϕ uma variável aleatória tal que sua densidade é dada por*

$$f(\phi) = \frac{1}{2\pi} \mathbf{I}_{[0, 2\pi[}(\phi)$$

Dizemos então que ϕ tem distribuição uniforme no intervalo $[0, 2\pi[$, e escrevemos $\phi \sim U(0, 2\pi)$.

Observa-se que a direção média de uma variável aleatória com distribuição uniforme não está bem definida.

Definição 3.2. *Seja ϕ uma variável aleatória tal que sua densidade é dada por*

$$f(\phi) = \frac{1}{2\pi I_0(\kappa)} \exp[\kappa \cos(\phi - \theta)] \mathbf{I}_{[0, 2\pi[}(\phi),$$

em que $\kappa \geq 0$, $\theta \in [0, 2\pi)$ e $I_0(\kappa)$ é a função de Bessel modificada do primeiro tipo de ordem 0.

Dizemos então que ϕ , dados θ e κ , tem distribuição de von Mises com ângulo médio θ e parâmetro de concentração κ , e escrevemos $\phi|\theta, \kappa \sim VM(\theta, \kappa)$.

Nota-se que, quando $\kappa = 0$, a distribuição de ϕ resultante nada mais é que a distribuição uniforme em $[0, 2\pi)$, qualquer que seja o valor de θ .

Iremos mostrar a seguir que, para uma variável aleatória ϕ que segue distribuição de von Mises com ângulo médio θ e parâmetro de concentração κ , a direção média, μ , é θ , e o comprimento médio resultante, ρ , vale $A(\kappa)$, com $A(\kappa) = I_1(\kappa)/I_0(\kappa)$. Para tanto, denotando

$2\pi I_0(\kappa)$ por \mathbb{C} , temos que:

$$\begin{aligned}\mathbb{E}[\cos \phi] &= \int_0^{2\pi} \frac{\cos \phi}{\mathbb{C}} \exp[\kappa \cos(\phi - \theta)] d\phi = \int_{0-\theta}^{2\pi-\theta} \frac{\cos(\psi + \theta)}{\mathbb{C}} \exp[\kappa \cos \psi] d\psi = \\ &= \int_0^{2\pi} \frac{1}{\mathbb{C}} (\cos \psi \cos \theta - \text{sen } \psi \text{sen } \theta) \exp[\kappa \cos \psi] d\psi = \mathbb{E}[\cos \psi] \cos \theta - \mathbb{E}[\text{sen } \psi] \text{sen } \theta, \quad (3.3.1)\end{aligned}$$

em que $\psi \sim VM(0, \kappa)$. Analogamente, mostra-se também que

$$\mathbb{E}[\text{sen } \phi] = \mathbb{E}[\text{sen } \psi] \cos \theta + \mathbb{E}[\cos \psi] \text{sen } \theta. \quad (3.3.2)$$

Temos também que $\int \frac{\text{sen } \psi}{\mathbb{C}} \exp[\kappa \cos \psi] d\psi = -\frac{\exp[\kappa \cos \psi]}{\kappa \mathbb{C}}$, de modo que

$$\mathbb{E}[\text{sen } \psi] = 0. \quad (3.3.3)$$

Dessa forma, temos, a partir de (3.3.1), (3.3.2) e (3.3.3), as seguintes relações:

$$\mathbb{E}[\cos \phi] = \mathbb{E}[\cos \psi] \cos \theta \quad (3.3.4)$$

$$\mathbb{E}[\text{sen } \phi] = \mathbb{E}[\cos \psi] \text{sen } \theta. \quad (3.3.5)$$

Também temos que

$$\mathbb{E}[\cos \psi] \equiv \int_0^{2\pi} \frac{\cos \psi}{\mathbb{C}} \exp[\kappa \cos \psi] d\psi = \frac{\int_0^{2\pi} \cos \psi \exp[\kappa \cos \psi] d\psi}{\frac{2\pi}{\frac{2\pi I_0(\kappa)}{2\pi}}} = \frac{I_1(\kappa)}{I_0(\kappa)} \equiv A(\kappa). \quad (3.3.6)$$

Assim, de (3.3.4), (3.3.5) e (3.3.6), sai que

$$\mathbb{E}[\cos \phi] = A(\kappa) \cos \theta$$

$$\mathbb{E}[\text{sen } \phi] = A(\kappa) \text{sen } \theta.$$

Como $A(\kappa) \geq 0$ (Mardia et al, [9]), temos que

$$A(\kappa) = \sqrt{\mathbb{E}^2[\cos \phi] + \mathbb{E}^2[\text{sen } \phi]} = \sqrt{\rho^2(\cos^2 \mu + \text{sen}^2 \mu)} = \rho,$$

e então, da relação (3.2.1), $\mu = \theta$.

A Figura 5 traz a função densidade de probabilidades para diferentes distribuições de von Mises, todas com ângulo médio π . Pode-se observar que quanto maior o parâmetro de concentração, menor a variabilidade.

Uma forma de se caracterizar a distribuição de von Mises é através da distribuição Normal Multivariada, como mostra o teorema a seguir:

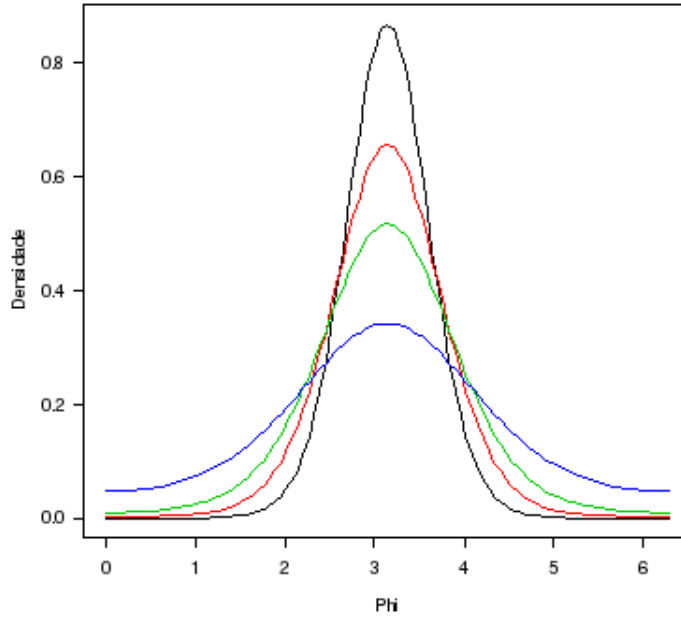


Figura 5: Densidade de variáveis Von Mises com ângulo médio π e parâmetro de precisão 5 (preto), 3 (vermelho), 2 (verde) e 1 (azul)

Teorema 3.1. *Seja (X, Y) um vetor aleatório tal que $(X, Y) \sim Normal(\mu, \Sigma)$, com*

$$\mu = \begin{bmatrix} m \\ d \end{bmatrix}$$

e

$$\Sigma = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}$$

Sejam $\phi \in [0, 2\pi)$ e r dados pela transformação (bijetora) a seguir:

$$re^{i\phi} = X + iY,$$

i.e., (r, ϕ) é a representação em coordenadas polares de (X, Y) . Então a distribuição de ϕ , condicional em $r = 1$, é de von Mises com parâmetros $\kappa = \sqrt{(m/\sigma^2)^2 + (d/\sigma^2)^2}$ e θ tal que $\sin \theta = \frac{d}{\kappa\sigma^2}$ e $\cos \theta = \frac{m}{\kappa\sigma^2}$, ou seja, $\kappa e^{i\theta} = \frac{m}{\sigma^2} + i\frac{d}{\sigma^2}$.

Demonstração. Primeiramente, observamos que o jacobiano da transformação dada no teorema vale

$$J = \begin{vmatrix} \cos \phi & -r \operatorname{sen} \phi \\ \operatorname{sen} \phi & r \cos \phi \end{vmatrix} = r.$$

Sejam κ e θ como no teorema. Temos que:

$$f_{(r,\phi)}(r, \phi) = r f_{(X,Y)}(r \cos \phi, r \operatorname{sen} \phi) \mathbf{I}_{[0,2\pi[}(\phi) = \frac{r}{2\pi\sigma^2} \exp\left\{-\frac{1}{2\sigma^2}[(r \cos \phi - m)^2 + (r \operatorname{sen} \phi - d)^2]\right\} \mathbf{I}_{[0,2\pi[}(\phi).$$

Logo,

$$\begin{aligned} f_{(\phi)}(\phi|r=1) &\propto \exp\left\{-\frac{1}{2\sigma^2}[\cos^2 \phi - 2m \cos \phi + m^2 + \operatorname{sen}^2 \phi - 2d \operatorname{sen} \phi + d^2]\right\} \mathbf{I}_{[0,2\pi[}(\phi) \propto \\ &\propto \exp\left[\frac{m}{\sigma^2} \cos \phi + \frac{d}{\sigma^2} \operatorname{sen} \phi\right] \mathbf{I}_{[0,2\pi[}(\phi). \end{aligned}$$

Se $m = d = 0$, temos imediatamente que a distribuição de ϕ , condicional em $r = 1$, é uniforme em $[0, 2\pi[$, isto é, $VM(\theta, 0)$, e o teorema está verificado. Se $m^2 + d^2 \neq 0$, e tomando κ e θ como no teorema:

$$f_{\phi|r=1}(\phi) \propto \exp\left\{\kappa\left[\frac{m}{\kappa\sigma^2} \cos \phi + \frac{d}{\kappa\sigma^2} \operatorname{sen} \phi\right]\right\} = \exp\{\kappa \cos(\phi - \theta)\}.$$

Logo, $\phi|r=1 \sim VM(\theta, \kappa)$. □

Geometricamente, esse teorema afirma que se tomamos, em \mathfrak{R}^2 , a distribuição conjunta de duas variáveis aleatórias independentes com distribuição normal univariada de mesma variância, então a distribuição condicional à circunferência de raio 1 centrada em $(0, 0)$ é de von Mises. Observa-se que essas curvas de nível são justamente circunferências. Uma versão deste teorema generalizada para distribuições de von Mises multivariadas encontra-se em Downs, [3]. Observa-se, ainda, que o resultado acima não é válido se $VAR[X] \neq VAR[Y]$, ou se a hipótese de independência entre X e Y não é válida. Nestes casos, a distribuição condicional resultante é a distribuição de Bingham (Mardia, [7]), que pode ser vista como uma generalização da distribuição de Von Mises.

3.4 Inferência

3.4.1 Distâncias e Funções de Perda no Círculo

Com o objetivo de se realizar estimativas pontuais acerca do parâmetro de interesse, deve-se estabelecer, do ponto de vista bayesiano, uma função de perda (ou, alternativamente, de utilidade), $L(a, \theta)$, que indica o modo como o decisor é punido para cada possível decisão a (estimativa pontual) e estado da natureza θ (valor do parâmetro) (DeGroot, [2]). Usualmente, para variáveis na reta, utiliza-se a perda quadrática, i.e., $L_q(a, \theta) = (a - \theta)^2$. A estimativa de Bayes de θ com relação à perda $L(a, \theta)$ é, então, o valor de a que minimiza o valor esperado da perda *a posteriori*, i.e., $\arg \min_a \mathbb{E}[L(a, \theta) | \mathbf{x}]$, em que \mathbf{x} é a amostra observada.

Nota-se, contudo, que a perda quadrática, bem como a perda absoluta ($L_a(a, \theta) = |a - \theta|$), não parecem razoáveis para uma variável circular. Isso ocorre devido à natureza periódica de um ângulo. Por exemplo, há uma diferença de 10^0 entre 350^0 e 0^0 , bem como entre 10^0 e 0^0 , ou seja, 350^0 e 10^0 são equidistantes de 0^0 . Contudo, tratando esses ângulos como números na reta, temos que 10 está mais próximo de 0 que 350 de 0.

Com a finalidade de solucionar tal problema, iremos a seguir definir duas medidas de distância no círculo e, a partir delas, duas funções de perda alternativas às usuais para variáveis na reta.

Sejam $\theta_a, \theta_b \in [0, 2\pi)$. Podemos definir as seguintes funções de $[0, 2\pi) \times [0, 2\pi)$ em \mathfrak{R}^+ : $d_{\cos}(\theta_a, \theta_b) = 1 - \cos(\theta_a - \theta_b)$ e $d_{abs*}(\theta_a, \theta_b) = \min\{|\theta_a - \theta_b|, 2\pi - |\theta_a - \theta_b|\}$.

Teorema 3.2. *As funções d_{\cos} e d_{abs*} são distâncias em $[0, 2\pi)$. Isto é, se $\theta_a, \theta_b, \theta_c \in [0, 2\pi)$, então:*

- (1) $d(\theta_a, \theta_b) = 0 \iff \theta_a = \theta_b$
 - (2) $d(\theta_a, \theta_b) = d(\theta_b, \theta_a)$
 - (3) $d(\theta_a, \theta_b) \leq d(\theta_a, \theta_c) + d(\theta_b, \theta_c)$
- (aqui $d = d_{\cos}$ ou $d = d_{abs*}$)

A demonstração deste fato encontra-se no Apêndice A.

Uma motivação para o uso destas distâncias para o círculo pode ser vista para o exemplo

numérico anterior: neste caso, temos que $d_{\cos}(350^0, 0^0) = 1 - \cos(-350^0) = 1 - \cos(10^0) = d_{\cos}(10^0, 0^0)$, ao passo que $d_e(350, 0) = \sqrt{(350)^2} \neq \sqrt{(10)^2} = d_e(10, 0)$. Em palavras, a distância d_{\cos} é a mesma para 350^0 e 0^0 e para 10^0 e 0^0 , contudo isso não é preservado na distância euclidiana (o que faz com que a perda quadrática torne-se, em geral, inadequada para dados circulares). Esta preservação também ocorre para a distância d_{abs^*} .

Desta maneira, é natural definir as seguintes funções de perda para o círculo: $L_{\cos}(a, \theta) = d_{\cos}(a, \theta)$ e $L_{abs^*}(a, \theta) = d_{abs^*}(a, \theta)$. A Figura 6 mostra a função $L_{\cos}(a, \theta)$ quando $\theta = \pi/4$.

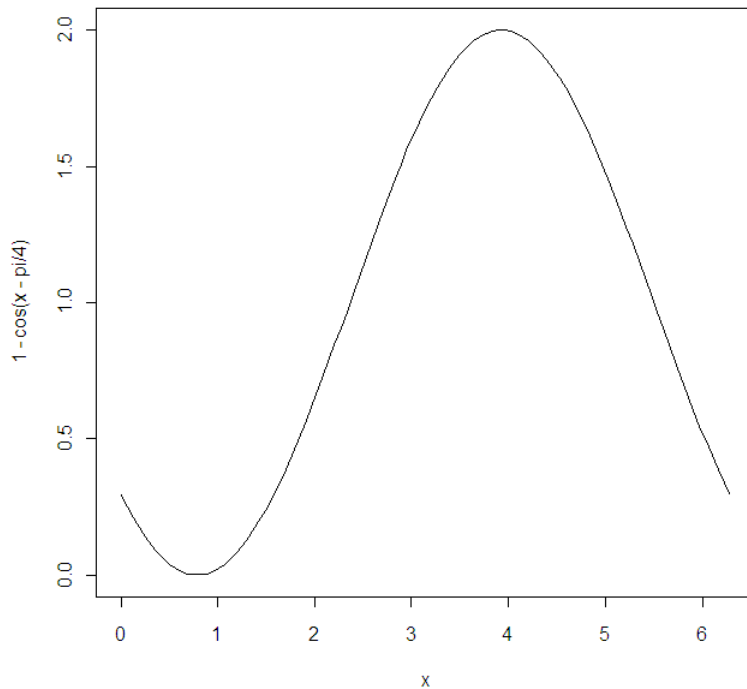


Figura 6: A função de perda $L_{\cos}(a, \theta)$, para $\theta = \pi/4$

O valor a que minimiza $\mathbb{E}[L_{\cos}(a, \theta)|\mathbf{x}]$ é justamente $\mu_{\theta|x}$, o ângulo médio da distribuição a *posteriori* de θ , dado x . De fato, sendo $\rho_{\theta|x}$ o comprimento médio resultante da distribuição

a posteriori de θ , dado x , temos:

$$\begin{aligned}
\mathbb{E}[L_{\cos}(a, \theta)|\mathbf{x}] &= 1 - \mathbb{E}[\cos(\theta - a)|\mathbf{x}] = 1 - \cos(a)\mathbb{E}[\cos\theta|\mathbf{x}] + \sin(a)\mathbb{E}[\sin\theta|\mathbf{x}] = \\
&= 1 + \rho_{\theta|x} \left[-\cos(a) \frac{\mathbb{E}[\cos\theta|\mathbf{x}]}{\rho_{\theta|x}} + \sin(a) \frac{\mathbb{E}[\sin\theta|\mathbf{x}]}{\rho_{\theta|x}} \right] = \\
&= 1 + \rho_{\theta|x} [-\cos(a) \cos \mu_{\theta|x} + \sin(a) \sin \mu_{\theta|x}] = \\
&= 1 - \rho_{\theta|x} \cos(a - \mu_{\theta|x}) = (**).
\end{aligned}$$

Como $\cos(a - \mu_{\theta|x})$ é maximizado quando $a = \mu_{\theta|x}$, esse é o valor que minimiza (**). Nota-se que, quando se utiliza esse valor de a para estimar θ , então quanto maior o valor do comprimento médio resultante a posteriori, $\rho_{\theta|x}$, menor a perda esperada.

Uma representação gráfica da função de perda $L_{abs*}(a, \theta)$, para $\theta = \pi/4$, encontra-se na Figura 7. Suponhamos que θ seja uma variável aleatória circular contínua e sejam $a_1 \in [0, \pi)$ e $a_2 \in [\pi, 2\pi)$ tais que $\mathbb{P}(a_1 < \theta < a_1 + \pi) = 1/2$ e $\mathbb{P}(a_2 - \pi < \theta < a_2) = 1/2$. Então, o valor de a que minimiza o valor esperado da perda a posteriori é $a = \arg \max_{x \in \{a_1, a_2\}} f_{\theta|\mathbf{x}}(x)$. Esse valor é muitas vezes chamado de mediana circular da distribuição a posteriori (Mardia et al, [9]). Observa-se que, em alguns casos, o mínimo de $L_{abs*}(a, \theta)$ pode ser atingido em vários pontos.

Observa-se ainda que a perda 0/1 continua tendo um papel importante na análise de dados circulares. A estimativa de Bayes com relação à perda 0/1 é a moda da distribuição a posteriori.

3.4.2 Resultados

O seguinte teorema é bastante útil para fins inferenciais de dados que seguem distribuição de von Mises:

Teorema 3.3. *Sejam $\phi_1, \dots, \phi_n | \theta, \kappa$ condicionalmente independentes e identicamente distribuídas $VM(\theta, \kappa)$. Temos então que $(C, S) \equiv (\sum_{i=1}^n \cos \phi_i, \sum_{i=1}^n \sin \phi_i)$ é suficiente para o par (θ, κ) .*

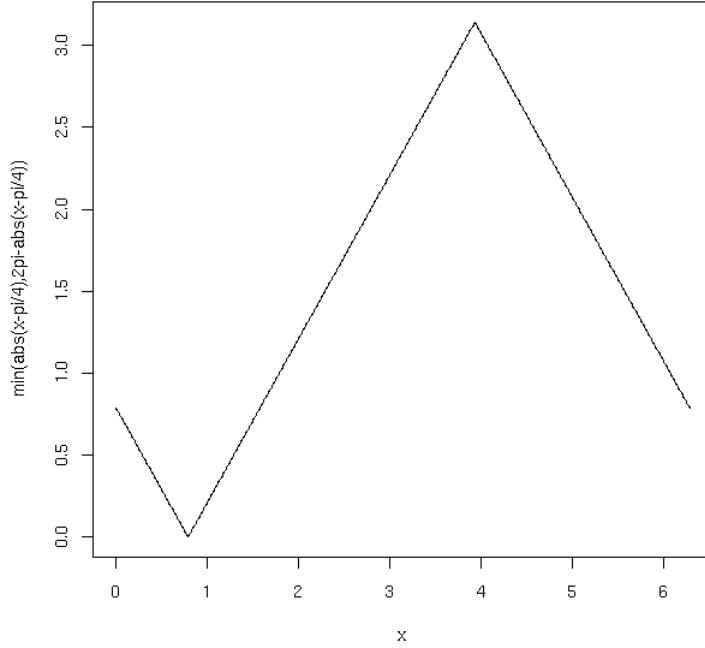


Figura 7: A função de perda $L_{abs^*}(a, \theta)$, para $\theta = \pi/4$

Demonstração. Temos que

$$\begin{aligned}
f(\phi, \dots, \phi_n | \theta, \kappa) &= \\
&= \prod_{i=1}^n \frac{1}{2\pi I_0(\kappa)} \exp[\kappa \cos(\phi_i - \theta)] \mathbf{I}_{[0, 2\pi]}(\phi_i) = \\
&= \left[\frac{1}{2\pi I_0(\kappa)} \right]^n \exp\left[\kappa \sum_{i=1}^n (\cos(\phi_i - \theta))\right] \prod_{i=1}^n (\mathbf{I}_{[0, 2\pi]}(\phi_i)) = \\
&= \left[\frac{1}{2\pi I_0(\kappa)} \right]^n \prod_{i=1}^n (\mathbf{I}_{[0, 2\pi]}(\phi_i)) \exp \left[\kappa \cos \theta \sum_{i=1}^n (\cos \phi_i) + \kappa \sin \theta \sum_{i=1}^n (\sin \phi_i) \right] \quad (3.4.1)
\end{aligned}$$

Assim, do critério da fatoração de Neyman (DeGroot, [2]), segue o resultado. \square

Sejam $\phi, \dots, \phi_n | \theta, \kappa$ condicionalmente independentes e identicamente distribuídos $VM(\theta, \kappa)$. Temos então que os estimadores de máxima verossimilhança de θ e κ são dados, respectivamente, por $\hat{\theta} = \bar{\phi}$ e $\hat{\kappa}$ tal que $A(\hat{\kappa}) = \bar{R}$, em que $A(\kappa) = I_1(\kappa)/I_0(\kappa)$. Isso pode ser verificado facilmente através do logaritmo da função de verossimilhança. De fato, usando (3.2.2) para

$\mu = \theta$ e (3.4.1):

$$\begin{aligned}\log L_{\phi_1, \dots, \phi_n}(\theta, \kappa) &= -n \log 2\pi - n \log I_0(\kappa) + \kappa \sum_{i=1}^n \cos(\phi_i - \theta) = \\ &= -n \log 2\pi - n \log I_0(\kappa) + n\kappa \bar{R} \cos(\bar{\phi} - \theta) = (*).\end{aligned}$$

Como a função cosseno atinge seu máximo em 0, obtemos $\hat{\theta} = \bar{\phi}$.

Agora, derivando-se (*) com relação à κ , temos:

$$\frac{\partial L_{\phi_1, \dots, \phi_n}(\theta, \kappa)}{\partial \kappa} = -n \frac{I'_0(\kappa)}{I_0(\kappa)} + n\bar{R} \cos(\bar{\phi} - \theta).$$

Como vale a relação $I'_0(\kappa) = I_1(\kappa)$, (Mardia et al, [9]) temos que $\hat{\kappa}$ é tal que $A(\hat{\kappa}) = \bar{R}$.

Observa-se que, mesmo quando κ é conhecido, o estimador de máxima verossimilhança de θ é $\bar{\phi}$.

Com a finalidade de se realizar inferência bayesiana sobre uma população que segue uma distribuição de von Mises, com κ conhecido, temos o seguinte teorema:

Teorema 3.4. *Suponhamos que, a priori, $\theta \sim VM(\theta_0, \kappa_0)$ e que $\phi_1, \dots, \phi_n | \theta, \kappa$ sejam condicionalmente independentes e identicamente distribuídos $VM(\theta, \kappa)$, e suponhamos κ conhecido. Sejam $r_1^* = \kappa_0 \cos \theta_0 + \kappa \sum_{i=1}^n \cos \phi_i$ e $r_2^* = \kappa_0 \sin \theta_0 + \kappa \sum_{i=1}^n \sin \phi_i$. Sejam ainda κ^* e θ^* tais que $\kappa^* e^{i\theta^*} = r_1^* + ir_2^*$. Temos então que $\theta | \phi_1, \dots, \phi_n, \kappa \sim VM(\theta^*, \kappa^*)$.*

Demonstração. Sejam $\theta, \phi_1, \dots, \phi_n$ como no teorema. Temos que:

$$\begin{aligned}f(\theta | \phi_1, \dots, \phi_n, \kappa) &\propto f(\phi_1, \dots, \phi_n | \theta, \kappa) f(\theta | \kappa) \propto \exp[\kappa \sum_{i=1}^n \cos(\phi_i - \theta)] \exp[\kappa_0 \cos(\theta - \theta_0)] = \\ &= \exp[\kappa \sum_{i=1}^n (\cos \phi_i \cos \theta + \sin \phi_i \sin \theta) + \kappa_0 (\cos \theta \cos \theta_0 + \sin \theta \sin \theta_0)] = \\ &= \exp[\cos \theta (\kappa_0 \cos \theta_0 + \kappa \sum_{i=1}^n \cos \phi_i) + \sin \theta (\kappa_0 \sin \theta_0 + \kappa \sum_{i=1}^n \sin \phi_i)] = \\ &= \exp[\cos(\theta) r_1^* + \sin(\theta) r_2^*] = (*).\end{aligned}$$

Se $\kappa^* = 0$, então $r_1^* = r_2^* = 0$, e o resultado vale, já que teremos $f(\theta | \phi_1, \dots, \phi_n, \kappa) \propto 1$ e $VM(\theta^*, \kappa^* = 0) \sim U(0, 2\pi)$.

Suponhamos então $\kappa^* \neq 0$. Teremos que:

$$(*) = \exp\left[\kappa^* \left(\cos \theta \frac{r_1^*}{\kappa^*} + \sin \theta \frac{r_2^*}{\kappa^*}\right)\right] = \exp[\kappa^* (\cos \theta \cos \theta^* + \sin \theta \sin \theta^*)] = \exp[\kappa^* \cos(\theta - \theta^*)],$$

de onde segue o resultado. □

Esta conjugação está descrita em uma versão para a distribuição de von Mises multidimensional em Mardia et al, [8]. Observa-se que quando $\kappa_0 = 0$, isto é, a distribuição *a priori* é uniforme, então, *a posteriori*, a direção média de θ é a direção média amostral.

Sejam $\phi_1, \dots, \phi_n | \theta, \kappa$ condicionalmente independentes e identicamente distribuídos $VM(\theta, \kappa)$, com κ conhecido. Suponhamos que, *a priori*, $\theta \sim VM(\theta_0, \kappa_0)$. Temos então, pelo Teorema (3.4), que $\theta | \phi_1, \dots, \phi_n, \kappa \sim VM(\theta^*, \kappa^*)$, com $\kappa^* e^{i\theta^*} = r_1^* + ir_2^*$, $r_1^* = \kappa_0 \cos \theta_0 + \kappa \sum_{i=1}^n \cos \phi_i$ e $r_2^* = \kappa_0 \sin \theta_0 + \kappa \sum_{i=1}^n \sin \phi_i$. Como o valor máximo da distribuição von Mises é atingido justamente no valor de seu ângulo médio, temos que as estimativas de Bayes com perda 0/1 e com perda $L_{\cos}(a, \theta)$ coincidem e são iguais a θ^* . Esse também é o valor da mediana circular da distribuição *a posteriori*.

Pode-se também calcular, para este caso, a distribuição preditiva de novas observações $\phi_{n+1}, \dots, \phi_{n+m}$ que, condicionalmente a θ e a κ , são independentes e identicamente distribuídos $VM(\theta, \kappa)$, sendo também condicionalmente independentes de ϕ_1, \dots, ϕ_n . Utilizando a mesma notação do Teorema (3.4), temos:

$$\begin{aligned} f(\phi_{n+1}, \dots, \phi_{n+m} | \phi_1, \dots, \phi_n, \kappa) &= \int_0^{2\pi} f(\phi_{n+1}, \dots, \phi_{n+m} | \theta, \kappa) f(\theta | \phi_1, \dots, \phi_n, \kappa) d\theta = \\ &= \int_0^{2\pi} \frac{1}{[2\pi I_0(\kappa)]^m} \exp\left\{\kappa \left[\sum_{i=n+1}^{n+m} \cos(\phi_i - \theta)\right]\right\} \frac{1}{2\pi I_0(\kappa^*)} \exp\{\kappa^* \cos(\theta - \theta^*)\} d\theta \\ &= \frac{1}{[2\pi I_0(\kappa)]^m 2\pi I_0(\kappa^*)} \int_0^{2\pi} \exp\left\{[\kappa \sum_{i=n+1}^{n+m} \cos \phi_i + \kappa^* \cos \theta^*] \cos \theta + [\kappa \sum_{i=n+1}^{n+m} \sin \phi_i + \kappa^* \sin \theta^*] \sin \theta\right\} d\theta = \\ &= \frac{I_0(R^*)}{[2\pi I_0(\kappa)]^m I_0(\kappa^*)}, \quad R^* = \left\{[\kappa \sum_{i=n+1}^{n+m} \cos \phi_i + \kappa^* \cos \theta^*]^2 + [\kappa \sum_{i=n+1}^{n+m} \sin \phi_i + \kappa^* \sin \theta^*]^2\right\}^{1/2}. \end{aligned}$$

A derivação desta distribuição preditiva, para $m = 1$, encontra-se em (Bagchi et al, [1]).

Para o caso geral em que $m \in \mathbb{N}$, a distribuição preditiva foi encontrada neste trabalho.

A forma da densidade preditiva não é de fácil tratamento analítico, de modo que o tratamento via simulação deste problema é mais simples.

Com a finalidade de analisarmos a sensibilidade da distribuição *a posteriori* frente aos dados, suponhamos que temos, *a priori*, $\theta \sim VM(\pi, \kappa_0)$ e que $\phi_1, \dots, \phi_n | \theta$ sejam condicionalmente independentes e identicamente distribuídos $VM(\theta, 1)$. Suponhamos que os valores observados são $\phi_1 = \dots = \phi_n = 0$. Então, utilizando o Teorema (3.4), vemos que, *a posteriori*, $\theta | \phi_1, \dots, \phi_n$ tem distribuição de von Mises com parâmetro de concentração $\kappa^* = |\kappa_0 - n|$. Assim, se $\kappa_0 = n$, teremos, *a posteriori*, uma distribuição uniforme. A Figura 8 traz as distribuições *a priori* e *a posteriori* para este caso.

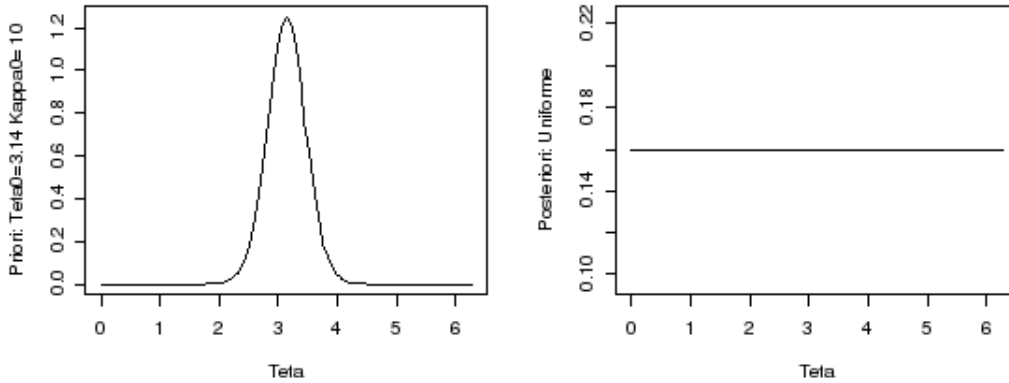


Figura 8: Distribuições *a priori* e *a posteriori* para o exemplo dado

Muitas vezes, tem-se interesse em testar a hipótese $H_0 : \theta = \theta_1$ contra $H_0 : \theta \neq \theta_1$, $\theta_1 \in [0, 2\pi)$. Podemos, sem perda de generalidade, restringir-nos ao caso em que $\theta_1 = 0$, uma vez que, caso isso não valha, podemos fazer um rotação nos dados e uma reparametrização da distribuição *a priori* de forma a cair nesse caso. Ainda no contexto em que κ é conhecido, o Teste da Razão de Verossimilhanças (TRV) (DeGroot, [2]) é aquele que rejeita a hipótese H_0 quando

$$\lambda(\phi_1, \dots, \phi_n) = \frac{\sup_{\theta \in \Theta_0} L(\theta; \phi_1, \dots, \phi_n)}{\sup_{\theta \in \Theta} L(\theta; \phi_1, \dots, \phi_n)} \leq C \iff \exp\left\{\sum_{i=1}^n \kappa \cos \phi_i + \sum_{i=1}^n \kappa \cos(\phi_i - \bar{\phi})\right\} \leq C \iff$$

$$\iff \exp\{\kappa n \bar{C} - \kappa n \bar{C} \cos \bar{\phi} - \kappa n \bar{S} \sin \bar{\phi}\} \leq C \iff \exp\{\kappa n \bar{C} - \kappa n \bar{R}\} \leq C.$$

Desta forma, o p-valor dado pelo TRV para uma amostra observada $\phi_1^{obs}, \dots, \phi_n^{obs}$ é dado por

$$\mathbb{P}(\bar{C} - \bar{R} \leq \bar{C}^{obs} - \bar{R}^{obs} | \theta = 0),$$

com

$$\bar{C}^{obs} = \frac{\sum_{i=1}^n \cos \phi_i^{obs}}{n},$$

$$\bar{S}^{obs} = \frac{\sum_{i=1}^n \sin \phi_i^{obs}}{n}$$

e

$$\bar{R}^{obs} = [(\bar{C}^{obs})^2 + (\bar{S}^{obs})^2]^{1/2}.$$

Como a distribuição de $\bar{C} - \bar{R} | \theta = 0$ não é conhecida, para que o p-valor seja calculado é necessário simular a distribuição desta estatística. Uma alternativa é utilizar o fato de que

$$-2 \log \lambda(\phi_1, \dots, \phi_n) = 2\kappa n (\bar{R} - \bar{C}) | \theta = 0 \approx \chi_1^2,$$

contudo para o uso de tal aproximação é necessária uma amostra grande.

Um segundo possível teste, talvez mais intuitivo, baseia-se no uso da distância circular d_{\cos} . A idéia deste teste é utilizar como p-valor em favor de H_0 a probabilidade do ângulo médio amostral, $\bar{\phi}$, estar mais distante (segundo d_{\cos}) de 0 que o ângulo médio amostral observado, $\bar{\phi}^{obs}$, sob H_0 , isto é,

$$\mathbb{P}(d_{\cos}(\bar{\phi}, 0) \geq d_{\cos}(\bar{\phi}^{obs}, 0) | \theta = 0) = \mathbb{P}(\cos \bar{\phi} \leq \cos \bar{\phi}^{obs} | \theta = 0).$$

Como a distribuição de $\bar{\phi}$ sob H_0 não é conhecida, para que o p-valor seja calculado, é necessário simular a distribuição desta estatística.

Observa-se também que $d_{\cos}(a, 0) \leq d_{\cos}(b, 0) \iff d_{abs^*}(a, 0) \leq d_{abs^*}(b, 0)$, de modo que se a distância utilizada no teste anterior fosse d_{abs^*} , e não d_{\cos} , o mesmo procedimento seria obtido.

Do ponto de vista bayesiano, tal hipótese pode ser testada utilizando-se o FBST (Pereira e Stern, [12]), que propõe como medida de evidência em favor de H_0 o e-valor, que, para a

hipótese dada, utilizando-se a distribuição *a priori* dada pelo Teorema 3.4, é

$$1 - 2 \min\{\mathbb{P}(0 < \theta \leq \theta^* | \phi_1, \dots, \phi_n, \kappa), 1 - \mathbb{P}(0 < \theta \leq \theta^* | \phi_1, \dots, \phi_n, \kappa)\}.$$

Pode-se observar que o cálculo do e-valor é mais simples que o do p-valor dado pelo TRV, uma vez que, para o seu cálculo, é necessário apenas um método para calcular a função de distribuição acumulada de variáveis aleatórias com distribuição de Von Mises. Funções que implementam tais métodos estão disponíveis em pacotes para o *R* como o *CircStats* e o *circular* (Jammalamadaka et al, [5]). Além disso, trata-se de um teste exato, uma vez que não é necessário recorrer a resultados assintóticos.

A seguir apresentamos três figuras que têm a finalidade de comparar os testes citados em alguns aspectos. Para que tal comparação fosse possível, utilizou-se como valor de corte para o FBST o valor que faria com que o tamanho do teste fosse dado por $\alpha = 5\%$.

A Figura 9 traz para cada um dos testes citados a região de aceitação (azul) e rejeição (vermelho) para cada possível amostra (ϕ_1, ϕ_2) de tamanho 2.

A Figura 10 traz para cada um dos testes citados a região de aceitação (azul) e rejeição (vermelho) para cada possível combinação (\bar{C}, \bar{S}) .

A Figura 11 traz a função poder de cada um dos testes citados para diferentes pontos.

O seguinte teorema, cuja demonstração encontra-se no Apêndice A, apresenta alguns comportamentos limites do e-valor para o caso estudado.

Teorema 3.5. *Suponhamos que, a priori, $\theta \sim VM(\theta_0, \kappa_0)$ e que $\phi_1, \dots, \phi_n, \dots | \theta, \kappa$ sejam condicionalmente independentes e identicamente distribuídos $VM(\theta, \kappa)$, e suponhamos κ conhecido. Sejam $r_{1n}^* = \kappa_0 \cos \theta_0 + \kappa \sum_{i=1}^n \cos \phi_i$ e $r_{2n}^* = \kappa_0 \sin \theta_0 + \kappa \sum_{i=1}^n \sin \phi_i$. Sejam ainda κ_n^* e θ_n^* tais que $\kappa_n^* e^{i\theta_n^*} = r_{1n}^* + ir_{2n}^*$. Consideremos a sequência de variáveis aleatórias $ev_n = 1 - 2 \min\{\mathbb{P}(0 < \theta \leq \theta_n^* | \phi_1, \dots, \phi_n, \kappa), 1 - \mathbb{P}(0 < \theta \leq \theta_n^* | \phi_1, \dots, \phi_n, \kappa)\} = 1 - 2 \min\{\mathbb{P}(0 < \theta_n \leq \theta_n^* | \phi_1, \dots, \phi_n, \kappa), 1 - \mathbb{P}(0 < \theta_n \leq \theta_n^* | \phi_1, \dots, \phi_n, \kappa)\}$, $\theta_n \sim VM(\theta_n^*, \kappa_n^*)$ (pelo teorema 3.4). Temos então que:*

- (1) $ev_n | \theta = \theta_0 \longrightarrow 0$ q.c., $\forall \theta_0 \in (0, 2\pi)$.
- (2) Se $\kappa_0 = 0$, $ev_n | \theta = 0 \sim Unif(0, 2\pi)$, $\forall n \in \mathbb{N}$.

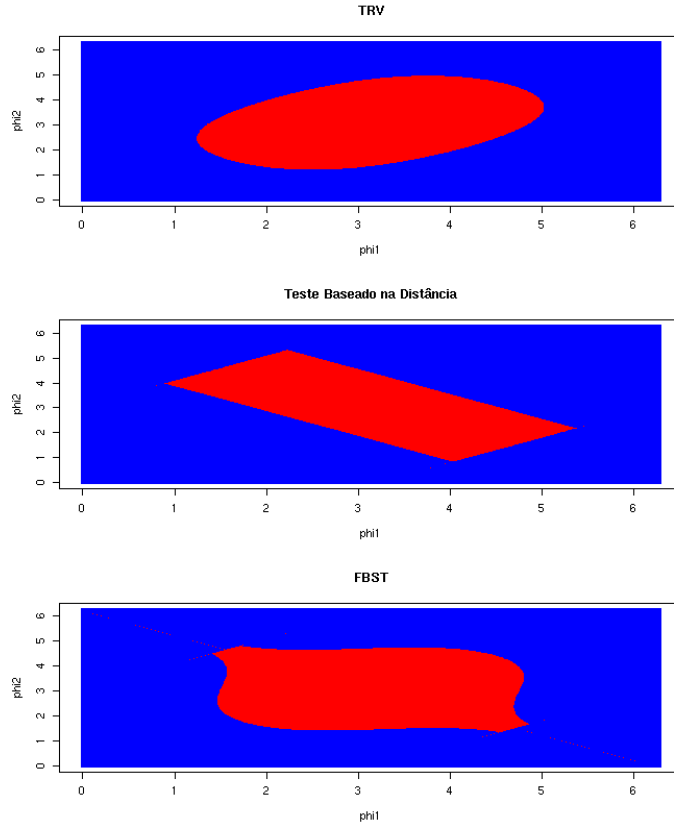


Figura 9: Região de aceitação (azul) e rejeição (vermelho) para os 3 testes. $n = 2$, $\kappa = 1$, $\alpha = 5\%$, distribuição *a priori* uniforme.

O teorema acima diz que o e-valor converge quase certamente a 0 quando a hipótese de que $\theta = 0$ não é verdadeira, e, quando a distribuição *a priori* é uniforme, tem distribuição uniforme se $\theta = 0$.

Em geral, o valor de κ não é conhecido. Uma forma de lidar com isso é através do teorema a seguir:

Teorema 3.6. *Sejam $\phi_1, \dots, \phi_n | \theta, \kappa$ sejam condicionalmente independentes e identicamente distribuídos $VM(\theta, \kappa)$. Suponhamos que, a priori, o par (θ, κ) seja distribuído segundo a densidade*

$$f(\theta, \kappa) \propto \frac{1}{I_0^c(\kappa)} \exp[\kappa R_0 \cos(\theta - \theta_0)] I_{\mathbb{R}_+^*}(\kappa) I_{[0, 2\pi)}(\theta).$$

Temos então que, a posteriori,

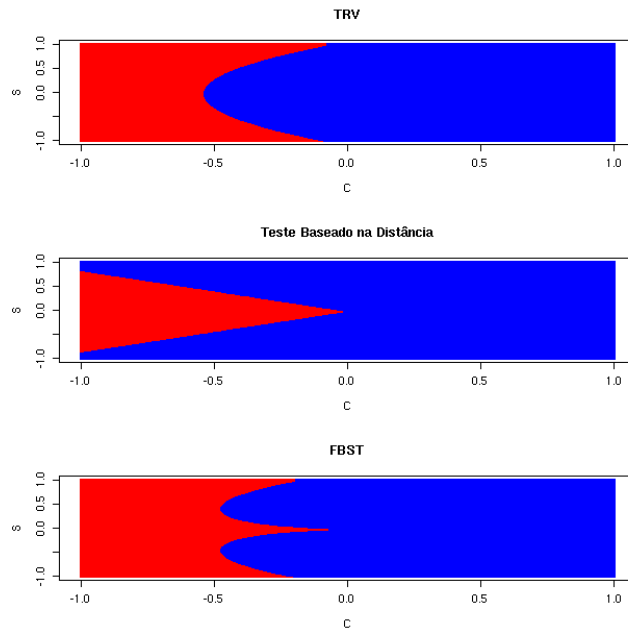


Figura 10: Região de aceitação (azul) e rejeição (vermelho) para os 3 testes. $n = 2$, $\kappa = 1$, $\alpha = 5\%$, distribuição *a priori* uniforme.

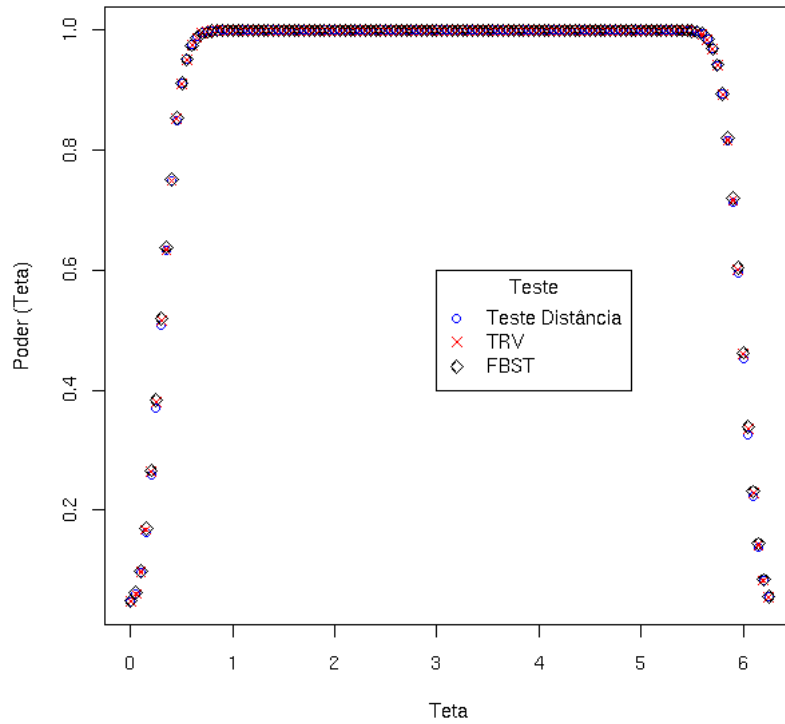


Figura 11: Função poder para os 3 testes. $n = 100$, $\kappa = 1$, $\alpha = 5\%$, distribuição *a priori* uniforme.

$$f(\theta, \kappa | \phi_1, \dots, \phi_n) \propto \frac{1}{I_0^{c^*}(\kappa)} \exp[\kappa R^* \cos(\theta - \theta^*)] I_{\mathbb{R}_+^*}(\kappa) I_{[0, 2\pi)}(\theta),$$

com $c^* = c + n$, e R^* e θ^* tais que $R^* e^{i\theta^*} = r_1^* + ir_2^*$, $r_1^* = R_0 \cos \theta_0 + \kappa \sum_{i=1}^n \cos \phi_i$ e $r_2^* = R_0 \sin \theta_0 + \kappa \sum_{i=1}^n \sin \phi_i$.

Demonstração. Sejam $\theta, \kappa, \phi_1, \dots, \phi_n$ como no teorema. Temos que:

$$\begin{aligned} f(\theta, \kappa | \phi_1, \dots, \phi_n) &\propto f(\phi_1, \dots, \phi_n | \theta, \kappa) f(\theta, \kappa) \propto \\ &\propto \frac{1}{I_0^n(\kappa)} \exp\left[\kappa \sum_{i=1}^n \cos(\phi_i - \theta)\right] \frac{1}{I_0^c(\kappa)} \exp[\kappa R_0 \cos(\theta - \theta_0)] I_{\mathbb{R}_+^*}(\kappa) I_{[0, 2\pi)}(\theta) = \\ &= \frac{1}{I_0^{c^*}(\kappa)} I_{\mathbb{R}_+^*}(\kappa) I_{[0, 2\pi)}(\theta) = \exp\left\{\kappa \left[(R_0 \cos \theta_0 + \sum_{i=1}^n \cos \phi_i) \cos \theta + (R_0 \sin \theta_0 + \sum_{i=1}^n \sin \phi_i) \sin \theta \right]\right\} = \\ &= \frac{1}{I_0^{c^*}(\kappa)} \exp[\kappa R^* \cos(\theta - \theta^*)] I_{\mathbb{R}_+^*}(\kappa) I_{[0, 2\pi)}(\theta). \end{aligned}$$

□

Uma extensão desse caso para mais dimensões encontra-se em (Nuñez et al., [10]). A conjugação descrita no Teorema (3.6) nada mais é que a conjugação natural (Paulino et al, [11]) da distribuição de Von Mises (com ambos os parâmetros desconhecidos), uma vez que esta pertence à família exponencial.

A forma desta distribuição conjugada não é de fácil tratamento analítico. Nuñez et al., [10], propõe uma maneira de simular valores da distribuição *a posteriori*, com base na técnica conhecida como *Sampling-Importance-Resampling* (SIR). Nota-se, contudo, que mesmo via simulação o problema da inferência sobre uma distribuição de Von Mises quando ambos os parâmetros são desconhecidos não é simples.

4 Conclusões

Pelo que foi estudado neste projeto, pode-se notar que, de fato, novas técnicas de análise estatística, bem como novos modelos probabilísticos, são necessários com a finalidade de

analisar dados circulares. Muitas vezes, tais técnicas apresentam uma dificuldade maior de implementação que as técnicas usuais para variáveis na reta.

Observa-se também que resta muito a ser feito, especialmente do ponto de vista bayesiano, para o problema de inferência em dados circulares.

Também nota-se que o cálculo do e-valor mostrou-se mais simples se comparado ao p-valor para o caso estudado, além de se tratar de um teste exato.

A Demonstrações adicionais

Neste apêndice iremos demonstrar os Teoremas (3.2) e (3.5).

A.1 Teorema (3.2)

Demonstração do Teorema (3.2). (1) e (2) são imediatos tanto para d_{\cos} quanto para d_{abs^*} . Iremos agora provar (3).

Seja d_e a distância euclidiana em \mathfrak{R}^2 , e sejam $l = (\cos \theta_l, \text{sen } \theta_l)$, $l = a, b, c$. Partindo da desigualdade triangular para a distância euclidiana, temos que:

$$\begin{aligned} d_e(a, b) &\leq d_e(a, c) + d_e(b, c) \Rightarrow \sqrt{(\cos \theta_a - \cos \theta_b)^2 + (\text{sen } \theta_a - \text{sen } \theta_b)^2} \leq \\ &\leq \sqrt{(\cos \theta_a - \cos \theta_c)^2 + (\text{sen } \theta_a - \text{sen } \theta_c)^2} + \sqrt{(\cos \theta_b - \cos \theta_c)^2 + (\text{sen } \theta_b - \text{sen } \theta_c)^2} \Rightarrow \\ \sqrt{1 - \cos(\theta_a - \theta_b)} &\leq \sqrt{1 - \cos(\theta_a - \theta_c)} + \sqrt{1 - \cos(\theta_b - \theta_c)} \Rightarrow \\ d_{\cos}(\theta_a, \theta_b) &\leq d_{\cos}(\theta_a, \theta_c) + d_{\cos}(\theta_b, \theta_c) \end{aligned}$$

Desta forma (3) é válido para d_{\cos} .

Para mostrar (3) para d_{abs^*} , dividiremos em alguns casos. Os casos não contemplados aqui são simétricos aos outros.

(1) $|\theta_a - \theta_b| < 2\pi - |\theta_a - \theta_b|$, $|\theta_a - \theta_c| < 2\pi - |\theta_a - \theta_c|$, $|\theta_b - \theta_c| < 2\pi - |\theta_b - \theta_c|$. Neste caso, a desigualdade triangular passa a ser $|\theta_a - \theta_b| \leq |\theta_a - \theta_c| + |\theta_b - \theta_c|$, que segue da desigualdade usual em \mathfrak{R} .

(2) $|\theta_a - \theta_b| < 2\pi - |\theta_a - \theta_b|$, $|\theta_a - \theta_c| < 2\pi - |\theta_a - \theta_c|$, $|\theta_b - \theta_c| \geq 2\pi - |\theta_b - \theta_c|$. Neste caso, a desigualdade triangular passa a ser $|\theta_a - \theta_b| \leq |\theta_a - \theta_c| + 2\pi - |\theta_b - \theta_c|$, que é equivalente a $|\theta_a - \theta_b| + |\theta_b - \theta_c| \leq |\theta_a - \theta_c| + 2\pi$. Mas, de fato, $|\theta_a - \theta_b| + |\theta_b - \theta_c| \leq |\theta_a - \theta_b| + |\theta_a - \theta_b| + |\theta_a - \theta_c| = 2|\theta_a - \theta_b| + |\theta_a - \theta_c| \leq 2\pi + |\theta_a - \theta_c|$, em que a última desigualdade segue do fato que $|\theta_a - \theta_b| < 2\pi - |\theta_a - \theta_b| \Rightarrow |\theta_a - \theta_b| < \pi$.

(3) $|\theta_a - \theta_b| < 2\pi - |\theta_a - \theta_b|$, $|\theta_a - \theta_c| \geq 2\pi - |\theta_a - \theta_c|$, $|\theta_b - \theta_c| \geq 2\pi - |\theta_b - \theta_c|$. Neste caso, a desigualdade triangular passa a ser $|\theta_a - \theta_b| \leq 2\pi - |\theta_a - \theta_c| + 2\pi - |\theta_b - \theta_c|$, que é equivalente a $|\theta_a - \theta_b| + |\theta_a - \theta_c| + |\theta_b - \theta_c| \leq 4\pi$. Para provar essa última desigualdade,

suponhamos, sem perda de generalidade, que $\theta_a = \min\{\theta_a, \theta_b, \theta_c\}$, e $\theta_c = \max\{\theta_a, \theta_b, \theta_c\}$. Temos então que $|\theta_a - \theta_b| + |\theta_a - \theta_c| + |\theta_b - \theta_c| \leq |0 - \theta_b| + |0 - 2\pi| + |\theta_b - 2\pi| = 4\pi$.

(4) $|\theta_a - \theta_b| \geq 2\pi - |\theta_a - \theta_b|$, $|\theta_a - \theta_c| < 2\pi - |\theta_a - \theta_c|$, $|\theta_b - \theta_c| < 2\pi - |\theta_b - \theta_c|$. Neste caso, a desigualdade triangular passa a ser $2\pi - |\theta_a - \theta_b| \leq |\theta_a - \theta_c| + |\theta_b - \theta_c|$. Para mostrar que ela vale, basta notar que $2\pi - |\theta_a - \theta_b| \leq |\theta_a - \theta_b| \leq |\theta_a - \theta_c| + |\theta_b - \theta_c|$.

(5) $|\theta_a - \theta_b| \geq 2\pi - |\theta_a - \theta_b|$, $|\theta_a - \theta_c| \geq 2\pi - |\theta_a - \theta_c|$, $|\theta_b - \theta_c| < 2\pi - |\theta_b - \theta_c|$. Neste caso, a desigualdade triangular passa a ser $2\pi - |\theta_a - \theta_b| \leq 2\pi - |\theta_a - \theta_c| + |\theta_b - \theta_c|$, que é equivalente a $|\theta_a - \theta_c| \leq |\theta_b - \theta_c| + |\theta_a - \theta_b|$, que segue da desigualdade usual em \mathfrak{R} .

(6) $|\theta_a - \theta_b| \geq 2\pi - |\theta_a - \theta_b|$, $|\theta_a - \theta_c| \geq 2\pi - |\theta_a - \theta_c|$, $|\theta_b - \theta_c| \geq 2\pi - |\theta_b - \theta_c|$. Nota-se que isso não pode ocorrer, pois $|\theta_a - \theta_b| \geq \pi$ e $|\theta_a - \theta_c| \geq \pi \Rightarrow |\theta_b - \theta_c| < \pi$.

□

A.2 Teorema (3.5)

Para a demonstração deste teorema, utilizaremos o seguinte lema, que pode ser encontrado em (Mardia et al, [9]):

Lema A.1. *Se ϕ_1, \dots, ϕ_n são independentes e identicamente distribuídas $VM(\theta, \kappa)$, então $\bar{\phi}|R = R_0 \sim VM(\theta, \kappa R_0)$, com $R = n\bar{R}$.*

Demonstração do Teorema (3.5). (1) Suponhamos, sem perda de generalidade, que $0 < \theta_0 < \pi$.

Não é difícil ver que $\kappa_n^*|\theta = \theta_0 \rightarrow \infty$ q.c., e também que $\theta_n^*|\theta = \theta_0 \rightarrow \theta_0$ q.c.. Seja $A = \{\omega \in [0, 2\pi)^\mathbb{N} : \kappa_n^*(\omega) \rightarrow \infty \wedge \theta_n^*(\omega) \rightarrow \theta_0\}$. Então $\mathbb{P}(A|\theta = \theta_0) = 1$.

Seja $\omega \in A$. Provaremos que $ev_n(\omega) \rightarrow 0$. Para tanto, fixemos $\epsilon > 0$ t.q. $0 < \theta_0 - \epsilon < \theta_0 + \epsilon < \pi$. Como $\omega \in A$, existe n_1 t.q. $n > n_1 \Rightarrow \theta_0 - \epsilon < \theta_n^*(\omega) < \theta_0 + \epsilon$. Além disso, $n > n_1 \Rightarrow ev_n(\omega) = 1 - 2\mathbb{P}(0 < \theta_n(\omega) \leq \theta_n^*(\omega))$. Notemos também que

$$\begin{aligned}\mathbb{P}(0 < \theta_n(\omega) \leq \theta_n^*(\omega)) &= \int_0^{\theta_n^*(\omega)} \frac{1}{2\pi I_0(\kappa_n^*(\omega))} \exp[\kappa_n^*(\omega) \cos(\theta - \theta_n^*(\omega))] d\theta \\ &= \int_0^{\theta_n^*(\omega)} \frac{1}{2\pi I_0(\kappa_n^*(\omega))} \exp[\kappa_n^*(\omega) \cos \theta] d\theta.\end{aligned}$$

Deste modo,

$$\begin{aligned}a_n &\equiv \int_0^{\theta_0 - \epsilon} \frac{1}{2\pi I_0(\kappa_n^*(\omega))} \exp[\kappa_n^*(\omega) \cos \theta] d\theta \leq \mathbb{P}(0 < \theta_n(\omega) \leq \theta_n^*(\omega)) \\ &\leq \int_0^{\theta_0 + \epsilon} \frac{1}{2\pi I_0(\kappa_n^*(\omega))} \exp[\kappa_n^*(\omega) \cos \theta] d\theta \equiv b_n.\end{aligned}$$

Provaremos que $\lim a_n = \lim b_n = 1/2$, de modo que $\lim \mathbb{P}(0 < \theta_n(\omega) \leq \theta_n^*(\omega)) = 1/2$ e, portanto, $\lim ev_n(\omega) = 0$.

Para tanto, fixemos $c \in (0, \pi)$. Temos que $f_n(\theta) \equiv \frac{1}{2\pi I_0(n)} \exp[n \cos \theta]$ é decrescente em $(0, \pi)$, pois $f'_n(\theta) = -\frac{\sin \theta}{2\pi I_0(n)} \exp[n \cos \theta] < 0$ quando $\theta \in (0, \pi)$. Assim, vale que $\int_c^\pi f_n(\theta) d\theta \leq f_n(c)(\pi - c) \rightarrow 0$. Desta forma, pela simetria de $f_n(\theta)$, $\int_0^c \frac{1}{2\pi I_0(n)} \exp[n \cos \theta] d\theta \rightarrow 1/2$, o que conclui a prova da parte (1), uma vez que $\kappa_n^*(\omega) \rightarrow \infty$ e $0 < \theta_0 - \epsilon, \theta_0 + \epsilon < \pi$.

(2) Seja $n \in \mathbb{N}$. Sejam $\kappa^* = \kappa_n^*$ e $\theta^* = \theta_n^*$. Do teorema 3.4, temos que $\kappa^* \cos \theta^* = n\kappa\bar{C}$ e $\kappa^* \sin \theta^* = n\kappa\bar{S}$, pois $\kappa_0 = 0$. Seja $x \in [0, 2\pi)$. Temos então que

$$\begin{aligned}\mathbb{P}(ev_n \leq x | \theta = 0) &= \mathbb{P}(\min\{\mathbb{P}(0 < \theta_n \leq \theta^*), 1 - \mathbb{P}(0 < \theta_n \leq \theta^*)\} \geq \frac{1-x}{2} | \theta = 0) \\ &= \mathbb{P}(\mathbb{P}(0 < \theta_n \leq \theta^*) \leq \frac{1+x}{2} | \theta = 0) - \mathbb{P}(\mathbb{P}(0 < \theta_n \leq \theta^*) \leq \frac{1-x}{2} | \theta = 0) = (*)\end{aligned}$$

Mas, utilizando o Lema (A.1) e o fato de que $\kappa_0 = 0 \Rightarrow \theta^* = \bar{\phi}$ e $\kappa^* = \kappa R$, temos que, para

$y \in (0, 1)$,

$$\begin{aligned}
& \mathbb{P}(\mathbb{P}(0 < \theta_n \leq \theta^*) \leq y | \theta = 0) = \\
& = \int_{\mathfrak{R}^+} \mathbb{P}(\mathbb{P}(0 < \theta_n \leq \theta^*) \leq y | \kappa^* = \kappa_1, \theta = 0) f_{\kappa^* | \theta = 0}(\kappa_1) d\kappa_1 = \\
& = \int_{\mathfrak{R}^+} \mathbb{P}\left(\int_0^{\theta^*} \frac{1}{2\pi I_0(\kappa^*)} \exp[\kappa^* \cos(\theta - \theta^*)] d\theta \leq y | \kappa^* = \kappa_1, \theta = 0\right) f_{\kappa^* | \theta = 0}(\kappa_1) d\kappa_1 = \\
& = \int_{\mathfrak{R}^+} \mathbb{P}\left(\int_0^{\theta^*} \frac{1}{2\pi I_0(\kappa_1)} \exp[\kappa_1 \cos \theta] d\theta \leq y | \kappa^* = \kappa_1, \theta = 0\right) f_{\kappa^* | \theta = 0}(\kappa_1) d\kappa_1 = \\
& = \int_{\mathfrak{R}^+} y f_{\kappa^*}(\kappa_1 | \theta = 0) d\kappa_1 = \\
& = y,
\end{aligned}$$

em que a penúltima igualdade segue do fato que $\theta^* | \kappa^* = \kappa_1, \theta = 0 \sim VM(0, \kappa_1)$ (pelo Lema (A.1)). Desta forma, $(*) = \frac{1+x}{2} - \frac{1-x}{2} = x$ e, portanto, $ev_n | \theta = 0 \sim Unif(0, 2\pi)$.

□

Referências

- [1] Bagchi, P. e Guttman, I. (1988). Theoretical considerations of the multivariate von Mises-Fisher distribution. *Journal of Applied Statistics* 15 149-169.
- [2] DeGroot, M. H. (1975). *Probability and Statistics*. Massachusetts: Addison-Wesley Publishing Company, Inc.
- [3] Downs, T. D. (1967). Some relationships between the normal and von Mises distributions. *Biometrika* 54 684-687.
- [4] Fisher, N. I. (1993). *Statistical Analysis of Circular Data*. Cambridge University Press.
- [5] Jammalamadaka S. R. e SenGupta A. (2001) *Topics in circular Statistics*. World Scientific.
- [6] Mardia, K. V. (1972). *Statistics of Directional Data*. London: Academic Press.
- [7] Mardia K. V. (1975). *Statistics of Directional Data*. *Journal of the Royal Statistical Society. Series B (Methodological)*. 37 349-393.
- [8] Mardia K. V. e El-Atoum S. A. M. (1976). Bayesian inference for the von Mises-Fisher distribution. *Biometrika* 63 203-206.
- [9] Mardia, K. V. e Jupp, P. E. (2000). *Directional Statistics*. Wiley.
- [10] Nuñez, G. A. e Peña, E. G. (2005). A Bayesian Analysis of Directional Data Using the von Mises-Fisher Distribution . *Communications in Statistics - Simulation and Computation* 34 989-999.
- [11] Paulino, C. D., Turkman, M. A. A. e Murteira, B. (2003). *Estatística Bayesiana*. Lisboa: Fundação Calouste Gulbenkian.
- [12] Pereira C. A. de B. e Stern J. M. (1999). Evidence and Credibility: a full Bayesian test of precise hypothesis. *Entropy*, 1, 99-110.

- [13] R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2005. URL <http://www.R-project.org>. ISBN 3-900051-07-0.